This work was originally published as: Daniel Gervais, Electronic Rights Management and Digital Identifier Systems - 4 Journal of Electronic Publishing (1999).

Electronic Rights Management and Digital Identifier Systems

The original version of this study was prepared for the first session of the Advisory Committee on Management of Copyright and Related Rights in Global Information Networks, held in Geneva on December 14th and 15th, 1998.

The new world of digital information requires a new way of providing access to that information — while keeping the copyright backbone. It might be technically easier to create a digital infrastructure without copyright: Just throw works up on the Internet, and let anyone get to them for any purposes. But such systems have been suggested and roundly rejected by those who create and own works of value. So we need to build an electronic infrastructure that works with copyright and takes advantage of the digital environment. This paper looks at the attempts to build part of that infrastructure — the electronic copyright-management system — and analyzes their progress and success.

Background

Before we can understand electronic copyright-management systems [1][#N1], we need to understand the concepts that underlie such systems, starting with "rights management" itself. Copyright-management systems are basically databases that contain information about content (works, discrete manifestations of works and related products) and, in most cases, the author and other current rights holders. That information is needed to support the process of authorizing the use of those works by others. A copyright-management system thus usually involves two basic modules, one for the identification of content and one for licensing or other rights transaction, such as a full assignment. In many cases, ancillary modules such as payment or accounts receivable are also considered part of the system, but the core of a copyright-management system is content and rights identification and a licensing tool.

A copyright-management system can be used by individual rights holders or by third parties who manage rights on behalf of others. A rights holder might use the system to track a repertory of works, manifestations, or products, or an organization representing a group of rights holders might use a copyright-management system to track each rights holder's rights and works. Such an organization might be a literary agent representing a number of writers, or, more commonly, a collective management organization such as an authors' society. (Most collective management organizations are members of the International Federation of Reproduction Rights Organisations or of the International Confederation of Societies of Authors and Composers.)

In a collective management organization, the mandate to authorize third parties may come directly from rights holders under a voluntary system, or it may follow from government regulations that require a non-voluntary license or that create a right to remuneration that must be managed collectively. [2][#N2] At the Copyright Clearance Center in the United States and some other similar

management organizations, rights holders set the price for each type of use of each piece of content. In the vast majority of cases, however, prices are contained in tariffs applicable to a class of content and/or users.

An annual fee, often referred to as a blanket license, is sometimes set by law, and covers an entire repertory of works. (A variation, the umbrella license, covers all content not specifically excluded by rights holders.) Repertory licenses are useful in cases where more precise management would be either too costly or simply impossible. Two good examples are music-performing rights (a radio station might purchase a blanket license to broadcast music), or corporate photocopying (a company might purchases a license to reproduce and distribute printed material within the company).

Another approach, the transactional fee, gives the user a license to use a specific work or manifestation for a defined purpose. For example, educational institutions in the United States that produce "coursepacks" (collections of photocopied material for students) must obtain prior authorization for each piece of content used, and that authorization covers only that coursepack; the use of music in advertising or, in most cases, to make a commercial recording also requires a transactional fee that covers only that specific use. In the transactional model, collective management organizations either grant a license based on agreed terms set in advance by the rights holder, or they act as an intermediary between the rights holder and the user to establish terms. [3][#N3]

Going Electronic

Using computers, rights holders can gain better access to large amounts of rights data, and today most use that electronic tool. The licensing function can also be made easier with computers, even though in practice progress has not been as rapid as with rights databases. Many so-called electronic copyright-management systems are in fact by and large done manually: E-mail may have replaced letters and faxes, but human intervention is necessary to process a license request. A fully automated licensing function that includes searchable online catalogues of prices, available content, and authorized uses, along with a "lights-out" licensing function (available 24 hours a day seven days a week) is still quite rare, but it is coming.

A system that would allow rights holders to maintain their rights data and automatically grant licenses to users could work without human intervention, which would have the benefit of keeping transaction costs low. The most successful of the systems deliver the content immediately, and they are being used by some journal and newspaper publishers and graphics houses.

Good systems function well both for repertory and transactional rights. For repertory rights, an electronic copyright-management system maintains a rights database that keeps track of both users and the works they are allowed access to, and often offers a license-renewal feature as well.

When one applies this concept of an electronic copyright-management system to practical rights licensing and trading applications, the enormous importance of such systems becomes readily apparent. To start with individual rights management, a full electronic copyright-management system (i.e., with both an automated rights information and licensing function) allows rights holders to maintain their rights data, to exchange it with other rights holders, and — more importantly — to process individual transactions from users. That is, licenses to use a specific work can be granted automatically to individual users. A corporation or an individual author or user can purchase the right to use an image, video clip, or song to republish it in a magazine article. A publishing house might purchase the right to reuse previously published material. Computer

software may be sold online with various types of licenses (e.g., use on a single computer, a site license, etc.). When performed online without human intervention, transaction costs are kept very low. Collective-management organizations represent a multitude of rights holders and deal with just as many users, and one of their principal advantages is the ability to put a large number of users in contact with a large repertory of content.

For transactional licenses, an electronic copyright-management system acts as a licensing "engine." There are various implementations of such systems that range in technical sophistication. In the least sophisticated, a user would mail, fax, or e-mail a license request to a collective management organization that would process it manually and return it to the user. In a slightly more automated environment, the organization would use an electronic works-and-rights database, but still process the license request manually. Another step up in the ladder of automation would be an internal computer-based licensing system that processes the request. With a full electronic copyright-management system, the user would search available content and rights online, submit a license request electronically (probably via the World Wide Web), and receive a response from the electronic copyright-management system without any human intervention. That last is, in my opinion, the only real electronic copyright-management system.

Ideally that true electronic copyright-management system can give the user not just the right, but the content itself. It is highly likely that future global information networks will allow users to search, select, download and at the same time clear the right to use a particular work, manifestation or product. Given the absence of global identification standards, however, an electronic copyright-management system may not be able to easily share works, rights and information relating thereto with another electronic copyright-management system.

As is rightly pointed out in material prepared by the International Bureau of the World Intellectual Property Organization (WIPO) for the important forum on collective administration held in Seville, Spain, in May 1997, until recently copyright management — when it was not performed by rights holders directly — was mostly known as "administration" and was divided by the medium of expression: music on records or tapes, films on 35mm or on video, photographs on paper or film, articles in books or periodicals. This is hardly surprising. All this has changed of course. Content is no longer media-dependent.

Most electronic copyright-management system are designed with interactive transmissions in mind. They may also apply to less precise forms of management, including the blanket licensing model mentioned above. In fact, the only way to have a true one-stop licensing shop is to combine all models, thus covering all types of rights and content. In other words, a one-stop shop must be able to clear use or reuse of all or part of a manifestation of any type of content, and may embrace also those rights, works and manifestations for which the rights holder prefers to use less precise forms of management, including blanket licenses.

Interactive Transmissions

The main area of application of electronic copyright-management systems is to clear rights for content transmitted on digital networks. The user may obtain both content and rights, or may already have the content and need the rights or additional rights. An electronic copyright-management system is the best, and perhaps the only way to clear rights in and monitor use of content sent through digital networks, since the usage itself can be monitored digitally on the networks subject to concerns about privacy. In fact, usage monitoring is already a key issue in the discussions between rights holders and access providers. Simply put, access providers agree to

return relevant market information to rights holders but want to limit their liability in case of user misuse.

A specific challenge of electronic copyright-management systems will be to allow managed access to material to be reused to create new products. That may be done on line, or could be done off line as when a user buys a password to use material contained on CD-ROM or DVD. Rights holders should have the option of delegating that task to a third party such as a collective management organization. The rights holder's decision to license through a third party should be based on factors such as cost, users' need for confidentiality of usage data, and efficiency. On the users' side, there is benefit in a central access point to identify information and clear rights.

An electronic copyright-management system must also allow users to obtain the rights that they want as precisely as possible, which means that ideally each item forming part of a given work or manifestation could be identified and dealt with separately. Traditionally, this type of rights clearance is handled manually by a rights and permissions or copyright department within a company. In some cases, part of the authorization may have been entrusted to a third party, such as a collective management organization. But the important point is that the process of rights management, which was exceptional compared to selling books and journals, is now moving to the forefront of the publishing industry because digital technology makes it much easier to reuse material. As evidence, just look at the number of meetings, seminars and articles in major newspapers about content available on the Web and on its reuse.

Legal Issues

The principal legal issues that need to be addressed in electronic copyright-management systems are ownership, rights to be conveyed, what the conveyance allows, and whose laws take precedence in case of a conflict.

- Who owns the rights?
  While the author usually owns the rights to a work, legal relationships like employment or work for hire may vest those rights in someone else. The issue is more complicated in the case of a manifestation like a film or play, where other rights holders like producers or performers may be involved. The electronic copyright-management system needs to know who owns the right to authorize the use of a work or manifestation in whole or part — and who may be entitled to a share of the royalties.

- Which rights are involved?
  Copyright is not a monolith. It comprises a number of different rights, and those rights have a separate existence in different parts of the world. We thus have a three-dimensional matrix, with a multitude of "rights" that can, in most cases, be separated territorially. [4][#N4]
  An inventory of the components of "copyright rights" is found in the Berne Convention and many national laws. There are two overarching categories: moral rights and economic rights. Within the former is the right of paternity or authorship and the right to oppose mutilation. In the latter category, the most important rights are the reproduction right, the right of communication to the public (which includes, according to Article 8 of the WIPO Copyright Treaty, the right to "make available"), and the right of adaptation. An electronic copyright-management system is concerned mainly with rights that can be licensed or traded on a routine basis, and therefore economic rights are better candidates for electronic rights management than moral rights.

- What rights are conveyed?
  A digital transmission implies making a copy, at least at the point of reception. Although some

argue that digital transmission involves the right of "distribution," a copy is not really distributed in the physical sense. In fact, whenever a protected work is accessed on a server and a user gets a copy, the right of reproduction rather than the right of distribution may be invoked. Certainly that is the position taken in the first of the Agreed Statements accompanying the WIPO Copyright Treaty (WCT):

> The reproduction right, as set out in Article 9 of the Berne Convention, and the exceptions permitted thereunder, fully apply in the digital environment, in particular to the use of works in digital form. It is understood that the storage of a protected work in digital form in an electronic medium constitutes a reproduction within the meaning of Article 9 of the Berne Convention."

Still, an open question remains about exceptions to the exclusive right of reproduction. As provided for in Article 9 of the Berne Convention, such exceptions, including "fair use" and "fair dealing," should have a limited scope whenever a commercial activity or any other wide-scale diffusion that interferes with the normal exploitation of the work is involved.

The right of communication to the public, which certainly applies to broadcasting, also applies to certain cases of interactive, on-demand transmissions. That issue certainly arises in regard to "push" technology, where information is sent to a user without his asking. Article 8 of the WCT says that the exclusive right of communication to the public includes "the making available to the public of their works in such a way that members of the public may access these works from a place and at a time individually chosen by them." That is a separate right, one that may be owned by a rights holder who does not also own the right of reproduction. If a particular use on the Web requires an authorization for both rights, two different clearance transactions may be necessary.

- Which country's laws take precedence?
  The traditional theories of emission (the law of the country of origin of the communication applies) and of reception (the law of the country of reception of the communication applies) are both very hard to transpose literally into a digital environment. One of the reasons is the multiplicity of countries that may qualify under either theory. When a user browsing the World Wide Web clicks to obtain remote content, he does not know whether that content comes directly from the site — and the host country — that the user is browsing. It may come from a mirror site in a third country. In that case, should we apply the fiction that the content came from the mother site? In other cases, sites or parts thereof are cached so that the content can be downloaded from a server closer to the user. Do we need a legal fiction to ignore the *actual* country of origin versus the *perceived* country of emission? With the country of emission approach, servers could be located in so-called "copyright havens." The reception theory seems simpler, and to a certain extent it is. The laws that take precedence are those of the country where the user is located. But that location is not always evident. As a resident of country A, I can use telephone lines to connect to the Internet in country B. To the system, I am located in country B. That problem may be evidentiary, but it matters.

  After a close reading of the Berne Convention, Andre Lucas recently proposed an amended version of the reception theory, which applies the law of the country where protection is required — the country for which protection is claimed (*lex loci delicti*). In most cases that would be the law of the country in which protection is claimed (*lex loci*), but not necessarily. Courts in a third country might be given jurisdiction by a contract between the litigants.

In addition, the question arises: How will the moral right apply? While the business of electronic

copyright clearly involves economic rights, it is impossible to ignore moral rights.

Good electronic copyright-management systems can handle ambiguity; they are not limited to Yes and No. Already, sophisticated electronic copyright-management systems are in use that can help protect moral rights two ways. First, since the system allows a contract between rights holder and user (with or without an intermediary), the parties may stipulate that alteration of the work is not allowed and/or that authorship must be recognized in a certain way. Second, rights holders can impose special conditions. For instance, with an electronic copyright-management system a photographer could insert language to restrict the use of her work to companies that she considers appropriate, specifically excluding tobacco companies. [5] [#N5]

Who Gets to Choose?

If the rights holder or service provider chooses the copyright-management system's environment, it is likely to be governed by laws of the country of emission. If the access provider makes the choice, it could be the country of emission, the country of reception, or a third country. If an electronic copyright-management system were used to manage rights at the user level (one could imagine a system integrated in a set-top box, for instance), the country of reception (or even the country in which the boxes are sold) could provide the environment. However, electronic copyright-management system modules may eventually be produced for a world market, granting rights more or less independent of any national law. The WIPO Copyright Treaty and other efforts by the World Intellectual Property Organization have increased harmony of national laws, and the gaps between national laws are closing, making that last scenario less fantastic.

But those gaps are not closing fast enough, and ambiguities remain. For instance, the Berne Convention, the Trade-Related Intellectual Property Rights Agreement, the WIPO Copyright Treaty, and the WIPO Performances and Phonograms Treaty impose limits on exceptions to common rules, but they *do* allow exceptions, and those exceptions vary considerably from one country to another. Depending on which law applies, an act may or may not require an authorization or may or may not be covered by a compulsory license or equitable remuneration scheme. It is still not clear what would happen if a French user were to download material from a U.S. site for educational uses in France: Would the U.S. Fair Use Guidelines [http://lcweb.loc.gov/copyright/fls/fl102.pdf] apply?

In addition, copyright is still negotiated and traded country by country and right by right. If I as an author have transferred the right to digitize and disseminate my work electronically to a publisher in, say, Hungary, what happens if a corporation in France downloads and copies my work from a site authorized by the publisher? Does that publisher have the right to authorize use in France? How does the Hungarian publisher even know that the corporation is in France? An electronic copyright-management system might include a function (e.g., a digital certificate or signature) that would check whether the user is located in a "valid" country, and that function could even include a digital-signature-based registration module that would confirm the mailing address of any user. [6] [#N6]

Standards Issues

Identifying what travels on digital networks is another essential part of a real-time electronic copyright-management system. The system must be able to precisely identify works, manifestations, and rights holders in order to secure authorizations from the right person, assign permissions, and then send payments to the rights holder. There are several competing standards

under consideration or in use today, many of them recognized by the International Organization for Standardization (ISO):

International Standard Work Code: This system, from the International Confederation of Societies of Authors and Composers, an umbrella organization representing a number of collective management organizations mainly in music, is being used for musical works and developed for literary works.

The codes are a "dumb" or "mute" numbers, in the sense that they do not in themselves contain any information. Unique for each object, the identification number is a key to a database where relevant information is contained. The music code (ISWC-T) consists of the letter T followed by a sequentially allocated ten-digit numeric code, the last digit of which is a "check digit" that allows the computer to validate the other nine digits. Numbers for the literary system (ISWC-L) will be similar.
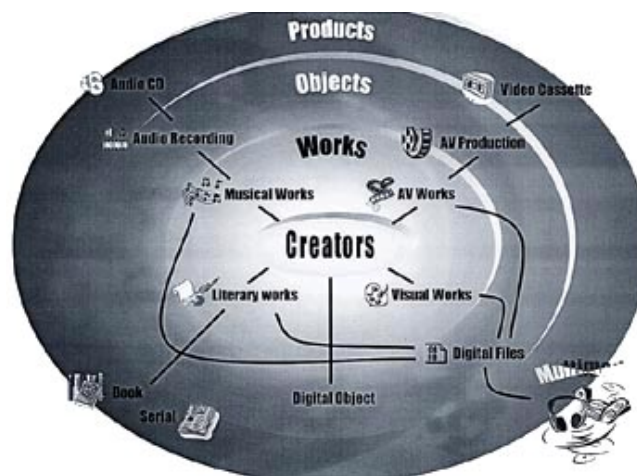
[/j/jep/images/3336451.0004.303-00000002.jpg]

The International Confederation of Societies of Authors and Composers' International Standard Work Code model:

Source: CISAC

World Congress, Berlin

September, 1998

International Standard Recording Code: According to the International Federation of the Phonographic Industry, the material traveling on electronic networks will not consist of "works" in a pure copyright sense, but rather of manifestations of works (also referred to as "digital objects"). Such manifestations might include a recording of a specific performance of a musical work (which, in the United States, may become a new work), or an HTML or PDF version of a scientific article published on the Web, including graphs and illustrations from various sources. Current IFPI identifiers for manifestations include the ISO-recognized International Standard Recording Code that identifies a musical recording (e.g., a track on a CD). Although it was adopted by ISO more than 10 years ago, less than 50% of recordings on the market have an embedded code.

International Standard Music Number: Another identifier in the music field is the ISO-recognized

International Standard Music Number, which is used for sheet music.

ISBN/ISSN: Books may be considered manifestations, although they are also finished commercial products. For over thirty years they have been identified using the International Standard Book Number (ISBN). The ISBN is composed of a one-digit "region" code, a publisher prefix, and then sequentially attributed numbers, followed by a check digit. Periodical publications are similarly identified at the title level by the International Standard Serial Number (ISSN), but that number applies to a periodical publication, not to the articles, graphs, charts, and images that it contains.

In the book trade, probably due to the absence of a specific publisher identifier, some people identify publishing houses by their ISBN prefix.

Barcodes: Rights holders may identify products like compact discs, computer software, or videos by putting a barcode on the distribution medium. The barcode may be based on the European Article Number or the Uniform Product Code. The barcode itself cannot be used in a purely digital environment, although the European Article Number and the Uniform Product Code might be adapted.

Publisher Item Identifier: Used in the publishing industry, the Publisher Item Identifier was developed in 1995 by an informal group of scientific and technical publishers: American Chemical Society, American Institute of Physics, American Physical Society, Elsevier Science and the Institute of Electrical and Electronics Engineers (IEEE). The Publisher Item Identifier is composed of seventeen alphanumeric characters that indicate publication type (whether it is a book or a journal), and other information depending on the type — such as the year of a serial publication. It contains no other intelligence, however, and is not linked to a central database. [7][#N7]

Serial and Book Item and Contribution Identifiers: [8][#N8] The Serial Item and Contribution Identifier (SICI) is a recognized standard used by serial publishers, subscription agents, and libraries, but no one has found a way to use it in the digital environment because it does not identify individual articles.

An expanded SICI and a new Book Item and Component Identifier (BICI) are now under development. They will be able to identify any part of a book or serial such as a chapter, an article, a foreword, an illustration, or a table.

The BICI is a flexible identification system with a fairly loose set of rules. The absence of firm rules here and in identifiers like the Digital Object Identifier reflect the amorphous and changing nature of the data to be identified, and the way in which it is stored, made available, and used or reused.

Compositeur, Auteur, Editeur Code: The CAE code is used by collective management organizations in the music field to identify those who create music and — more recently — other forms of information. Created in 1992 by the International Confederation of Societies of Authors and Composers, the code has been superseded by the IP number to identify "Interested Parties" to a work — a full range of rights holders. The format of the number itself did not change, and previously allocated CAE codes were converted into IP numbers. As with the some other identifiers, the numbers convey no meaning.

At present, use of and access to the IP database is restricted to confederation members. If it is made available, it could lead to a standard identifier for people by all copyright industries.

Digital Object Identifier: The DOI is not an identifier *per se*, but it offers both a structure for an identifier and a persistent routing system to a database containing relevant information.

Launched by the Association of American Publishers [http://www.publishers.org] in conjunction with the Corporation for National Research Initiatives [http://www.cnri.reston.va.us/about_cnri.html] at the 1997 Frankfurt Book Fair, the DOI was designed to "provide persistent and reliable identification of digital objects via a proven technology — the CNRI Handle System® [http://www.handle.net/index.html] — and an efficient administration system to link customers with publishers, facilitate electronic commerce, and enable automated copyright management systems." The CNRI Handle System is a distributed computer system that stores names of digital items and can quickly find the information necessary to locate and access the items.

The DOI is thus mainly two things: an identification *system*, potentially applicable to any and all categories of works and manifestations (even though at present its beta users are mostly book and journal publishers), and a central directory or database which, when queried using a DOI number, will route the user to the appropriate source of information.

The DOI number is composed of three parts:
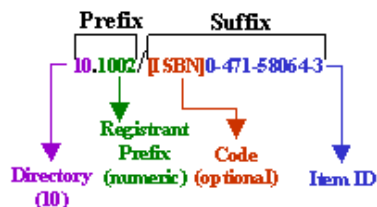
[/j/jep/images/3336451.0004.303-00000003.gif]

Source: The International DOI Foundation [http://www.doi.org/DOI-Found-Recruit.html]

The DOI is very flexible, given that rights holders or other persons using it as an identifier can use any suffix, including other existing identifiers (the ISBN in the example given above).

The DOI is functionally similar to a Uniform Resource Locator (URL) in that a user can click on it and go directly to the DOI Directory, which in turn seamlessly reroutes the user to the source of information corresponding to that DOI. Unlike a URL, the DOI can easily be rerouted. A rights holder who purchases rights to a work from another rights holder can update the Directory information to ensure that future clicks are routed to its system.

However, the DOI Foundation is still grappling with the issue of which digital objects the DOI should identify. [9][#N9] The creative communities as well as some more traditional copyright industries see the point of departure as the creative work or its manifestation. They see the initial work as the "core" to be identified, acknowledging that it may have digital "versions". Even from that viewpoint, however, the task is difficult, given that there is no uniform identification system for those works and manifestations and "no widely accepted data model defining all creative and publishing acts, necessary in placing [creations] in a digital world." [10][#N10] If the original works are identified using DOIs, should the various "physical" manifestations receive DOIs? What about products such as books, journals, articles and abstracts? The information industry, on the other hand, starts with digital objects that can be traded and has no need or desire to go "upstream" back to the original "work".

The conclusion drawn by the DOI Foundation is that no single identifier is capable of serving all

purposes. This is not fatal, however, because the DOI is not "just" an identifier. Rather, it is a structure in which other identifiers can be used to create a new identifier. With this structure in place, it is likely that the DOI and interested parties will be able to also offer an electronic copyright-management system solution, at least for print publications (in paper or digital form), probably late in 1999.

International Standard Audiovisual Number: The International Standard Audiovisual Number is a joint development of the International Confederation of Societies of Authors and Composers, the International Federation of Film Producers Associations, and the Association de Gestion Internationale Collective des Oeuvres Audiovisuelles. The audiovisual number has reached the level of "committee draft" within the International Standards Organization, and has been submitted to national ISO committees. [11][#N11]The proposed identifier is a sixteen-digit dumb number that may be used to identify audiovisual works of all kinds. [12][#N12]It is an identification number without any legal implication or meaning and has no *prima facie* evidence value as regards the copyright status or ownership of the work. It does not identify rights owners, even though it will be a tool used by people concerned with rights management as well as by many people interested in precise identification of audiovisual works. In other words, the number is a mere pointer to a database where information necessary for the identification of content is maintained.

The proposal is to affix the number onto the work — on masters and copies, whether in analog or digital format, on packaging, contacts, etc. The system will be administered by an *ad-hoc*, non-profit-making, international agency. The system and the information in the identification database will be open to any interested user. A fee will be charged to access the database. Many collective management organizations active in the audiovisual field plan to use the number as a key feature of the International Database on Audiovisual Works, a database of rights ownership in audiovisual works to be used for collective rights-management purposes.

Persistent Uniform Resource Identifiers: There are various proposals to upgrade the standard Internet Uniform Resource Locators (URLs). The problem is that when a digital resource moves from one "page" or file on a server to another, or from one server to another, the URL also changes. A user who enters the original URL in the browser gets the infamous "error 404" message, meaning that the resource is no longer available at that address.

The Internet Engineering Task Force, working with the Corporation for National Research Initiatives, developed the Uniform Resource Name (URN) [http://www.w3.org/Addressing/] . Like the URL, it is a uniform resource identifier, but one with an "institutional commitment to persistence." In other words, the institution that issued the URN (and presumably also put the resource on the Web) commits to maintain the validity of the URN, if necessary by posting a page that redirects users to the resource's current address.

One implementation of URNs is the Persistent Uniform Resource Locator (PURL). PURLs are in fact URLs that point to a server that can be updated (a system not unlike the DOI directory). "Instead of pointing directly to the location of an Internet resource, a PURL points to an intermediate resolution service. The PURL resolution service associates the PURL with the actual URL and returns that URL to the client. The client can then complete the URL transaction in the normal fashion. In Web parlance, this is a standard HTTP "redirect." [13][#N13]

Metadata Issues

Until and unless a single global identification system can be agreed upon, electronic copyright-management systems must be able to function in a multi-code environment. And that means that

information about the information — metadata — must be made available in a usable format.

While there are existing standards for bibliographic metadata that go back many decades, the situation is less clear in other sectors. In the audiovisual sector, there are databases that contain information like film credits, but currently there are no worldwide standards. Following are some of the efforts to develop metadata standards:

The Dublin Core: [http://purl.oclc.org/metadata/dublin_core/] The Dublin Core is an attempt to identify the "core" elements of metadata that are needed to satisfy the needs of all those involved in the exchange of or commerce in electronic-information resources. It was developed over a three-year period at workshops in which "experts from the library world, the networking and digital library research communities, and a variety of content specialties" participated. [14][#N14] This "core" was named after the city in Ohio in which the first meeting was held.

Originally the Dublin Core contained fifteen core elements: Title, Subject, Description, Creator (or primary contributor), Contributor, Publisher, Date, Type, Format, Identifier, Source (previous resource), Language, Relation (to another resource), Coverage (geographical or temporal) and Rights. In further meetings, other elements were added including the concept of a sub-element, which is used to qualify an element (for example, "date" can refer to a date of publication, or of a revision); a scheme, a label used to identify the method followed to identify the data (e.g., Dewey or MARC); and the language in which the metadata is entered, as opposed to the language of the resource itself.

A number of other groups are working on standards that could have a direct impact [http://www.w3.org/Metadata/Activity.html] on the future of the Dublin Core. While those standards are not for metadata *per se*, they affect the way metadata is coded, transmitted, used, retrieved, and accessed. For example, the World Wide Web Consortium is developing new markup languages and a new language for representing metadata in XML, the markup language designed to replace HTML. [15][#N15]

US MARC: A well-known public repository of metadata in the United States is the machine-readable cataloging-records database (known as US MARC). The US MARC formats are standards for the representation and communication of bibliographic and related information in machine-readable form. The US MARC was developed by the Library of Congress, the Canadian National Library, and the American Library Association along with the Australian National Library, Online Computer Library Center, the Music Library Association, and the Special Libraries Association. The US MARC database [http://lcweb.loc.gov/marc/] contains approximately seven million records of publisher titles. It can be searched online.

A US MARC record contains three elements: the record structure, the content designation, and the data content of the record. The US MARC benefits from the fact that it already applies to a vast number of titles. The question is whether (and how) it could be extended to apply to other types of content.

INDECS Project: A new project, Interoperability of Data in E-Commerce Systems, or INDECS, was launched in November 1998, funded by the European Commission. It is worth mentioning here because it seems to have the support of many international representative organizations. In addition, the project partners have significant experience working with identification and metadata issues in a copyright environment, and understand the difficulties encountered by the various ongoing standardization efforts. INDECS posits the existence of several identification schemes and, rather than try to achieve worldwide harmonization, is aiming for interoperability.

According to the Project description submitted to the European Commission, INDECS will try to deliver, within a fifteen-month timeline,

- A directory of participants, which should help provide a "translation" among existing or proposed identifiers;

- A directory of such identifiers; and

- An agreement on metadata formats.

At the heart of the project will be an effort to achieve agreement on a proposed standard for descriptive and rights metadata. Such a standard would provide a means of integrating data from the different rights-based sectors. To produce those deliverables, INDECS will also need to achieve agreement on a data dictionary to address the numerous vocabulary problems that stem from the participation of people and companies with widely different backgrounds and practices.

Stanford Digital Library Metadata Architecture:
[http://dbpubs.stanford.edu:8091/diglib/pub/delos.html] Part of the Stanford Digital Library Project, the Stanford Digital Library Metadata Architecture is another interesting effort to standardize metadata. The architecture does not contain a formal metadata proposal; it is a metadata "service layer" within the digital library that could interface with several repositories or databases of metadata (such as a Dublin Core repository, or the US MARC records) and external search engines, and would be able to translate different models with a view to achieving interoperability.

BIBLINK/NEDLIB: BIBLINK is not a metadata definition project as such, but rather a project that aims to establish a relationship and encoding model between national bibliographic agencies and publishers of electronic material, in order to establish authoritative bibliographic information that will benefit both sectors. It is intended to deliver an interactive demonstration system that will enable publishers of electronic documents to input and transmit an agreed minimum level of data describing the documents to national bibliographic services. [http://hosted.ukoln.ac.uk/biblink/] BIBLINK is funded by the European Commission.

The Networked European Deposit Library (NEDLIB) is sponsored by a group of European national libraries and particularly the National Library of the Netherlands. [16][#N16] It started where BIBLINK ended. Launched in January 1998 and funded by the European Commission, NEDLIB is not a metadata project. Its chief aim is to "construct the basic infrastructure upon which a networked European deposit library can be built."

Further work on NEDLIB might provide useful guidance on the use of metadata in transactions between librarians and publishers.

While the music and audiovisual fields have taken some steps to standardize metadata, there is no standard for data concerning rights ownership, licensing and trading. Thus, while the metadata may be used to identify a particular piece of content, it may not be sufficient or even useful for electronic-commerce transactions. If, for example, the rights holder is not the "original" rights holder indicated in the bibliographic metadata, the data could do more harm than good.

Technology Issues

Protection of rights-management information requires a synergy between law and technology, and there are several projects that are looking at that interface. Some important issues include:

The Protection of Electronic Copyright-Management Systems: Electronic copyright-management

systems themselves need protection. They need standard identification and delivery formats and tools to work automatically. With the growing use of digital networks to access protected content, it is highly likely that rights holders will invest heavily in identifying and permanently marking digital works. Worldwide implementation of the WIPO Copyright Treaty and the WIPO Performances and Phonograms Treaty should ensure that rights-management data are not deliberately altered.

Privacy and Confidentiality Issues: [17][#N17] Two of the questions most often asked by users are

1. As a private individual, can I browse/read/watch/listen without giving my identity (and then receiving mail and telephone solicitations, etc.)?

2. As a corporate user (say a pharmaceutical corporation), can I download this scientific article without the whole world knowing that I need this for my R & D efforts?

There may be legal grounds on which to base a claim for privacy or confidentiality in accessing protected content. In fact, academic authors in the United States have argued that the Constitution protects a right to read anonymously. [18][#N18] In many European countries, private data are protected and may be used only within strict guidelines.

An electronic copyright-management system does not in and by itself protect privacy, but it is probably the best tool to do so. If the rules under which the electronic copyright-management system operates are correctly designed, the system would return to rights holders aggregated information on use of his/her works. For example, the system could say that clearance was granted to use "Scientific Article X" to "11 pharmaceutical companies in the last month", or that "2,345 users in this part of Chicago" downloaded a given musical work. The rights holder thus gets market data without violating anyone's confidentiality or privacy. Even now the Copyright Clearance Center in the U.S. does not report to rights holders which articles from medical or scientific journals are used by individual users (e.g., pharmaceutical companies). It only tells rights holders how often a work was used by, say, the pharmaceutical industry as a whole. Most collective management organizations aggregate information in this way and this is perhaps a function whose value has thus far been underestimated by users.

A related issue is how to identify individual digital copies (which presumably have been sold to a specific user), without creating a risk to privacy or confidentiality. If indeed individual copies are identified, using a watermark containing a transaction code for instance, a viable solution could be to number individual copies, without including data identifying the user who "ordered" the copy in question. Copy numbers could be linked, in a secure database, to the individual users. Should there be a good reason to make the link between the copy number and the user — for instance, under court order — that link could be made. The role of trusted third parties acting as aggregators of usage data might be especially important to users. An aggregator or collective management organization using an electronic copyright-management system could thus maintain the confidentiality of the link (if any) between a given copy delivered on-line and a specific user. The content owner would receive with the payment for use of his works a report on the number of uses, possibly with an indication of the type of users concerned, but no information about individual users. Without this type of confidentiality guarantee, it may be very difficult for electronic copyright commerce to prosper. In other words, properly tuned electronic copyright-management systems that aggregate data so as to protect privacy and confidentiality are probably essential ingredients of the success of electronic copyright commerce.

Electronic Copyright Commerce

Electronic Copyright Commerce includes at least the two following types of transactions:

1. Online licensing (without content). (The user already has the content and wishes to perform a restricted act which requires an authorization); and

2. Online purchase of content protected by copyright, with the necessary rights clearance to allow the user to perform a restricted act.

In both cases, the intellectual-property aspect of the transaction is a licensing transaction, [19][#N19] and a license is a contract. In fact, contracts govern electronic commerce — within the parameters of copyright principles and other public policy ("*ordre public*") rules.

An electronic copyright-management system should allow users to access material in such a way that they know what they are buying. For example, a viable commercial model might allow users to freely browse part of the content — a few pages of an article, 30 seconds of music, a film preview — or view the work at a reduced graphic resolution. Full access would require payment. The right to commercially reuse the material would require a higher level of clearance.

The challenge of electronic copyright-management systems is to make all this possible in a user-friendly way. Content providers themselves may manage the system, but in most cases intermediaries will probably be involved. They will act as central access points, performing or allowing a user to perform a discovery function of available material from several content providers. Today's intermediaries, such as subscription agents and libraries, may not be the organizations that will provide electronic copyright-management systems. Such systems demand a level of economic and legal complexity and technological privacy guarantees that some intermediaries will not be able to provide. The best agents for rights management may be rights management agencies; in particular those with experience in the transactional model. Several of today's collective management organizations also might move into that field.

Existing Systems

A full electronic copyright-management system includes a work and a rights database, and a licensing function. Ideally, at least for certain types of content, it also offers a link to the content itself, thus allowing the user to clear the rights and obtain the content in a single set of transactions.

There are a number of working electronic copyright-management systems in operation today, even though the field is still fairly new — so new that a 1996 report concluded that the systems were still nascent, but today there are real products. The examples below include only those systems that bring together the works of several rights holders, not those run by and for a single rights holder.

ALCS/ByLine: The UK-based Author Licensing and Collecting Society has developed an electronic copyright-management system solution for online syndication of newspaper and other articles [formerly http://www.universalbyline.com/scoop.html]. Combining copyright management with online delivery, the system allows users to search the ByLine database and download articles for republication, with authority from the author of the material. The system is automated and articles are individually priced.

Copyright Clearance Center: The US-based CCC [http://www.copyright.com/] offers an electronic copyright-management system that is available on the World Wide Web, and that allows rights

holders to set their prices, establish acceptable uses, and view their accounts directly. The service is offered for universities interested in clearing photocopied coursepacks and electronic course content, and for general photocopying permissions. CCC also offers on-line licensing of specific titles for commercial republication. [http://www.copyright.com]

As a member of the International Federation of Reproduction Rights Organisations, [http://www.ifrro.org/] the CCC has an interface to the copyright-management systems (some of them off line) of IFRRO members such as the UK Copyright Licensing Agency's Rapid Clearance Service [http://www.cla.co.uk/support/clarcs/index.html] (CLARCS) and the Australian Copyright Agency Limited's Copyright Xpress. [Editors note: link to Copyright Xpress removed August 2001 because it seems that the Australian Copyright Agency no longer offers that service.]

Mira: [http://www.mira.com] The Media Image Resource Alliance is an online digital-stock agency. Users can browse, download, and clear rights to use professional-quality images. The entire licensing function and access to content is done automatically online via an electronic copyright-management system. Mira was created by the American Society of Media Photographers, the Copyright Clearance Center, and Applied Graphics Technology. Photographers and other rights holders provide images directly to Mira, and set prices and conditions for use.

The Way Forward: Interoperability

It is not realistic to see the world as a series of separate, parallel, watertight delivery systems. Interoperability is Ariadne's thread in the labyrinth of electronic-copyright commerce.

But history is against us. WIPO's Film Register Treaty, which would have allowed rights holders and users alike to find rights-ownership information easily, and which could have facilitated trade and enhanced transparency in that industry, did not receive the support of major film studios. Perhaps it came too soon, and certainly its focus was not digital, but the argument against it was that rights holders did not need a metadata database due to their reliance on their own data and worldwide distribution networks. Concerns were also raised about possible use of the registry by pirates.

Today large entertainment, publishing and other conglomerates have developed or are currently developing proprietary solutions for identification, rights management, and delivery of digital content. While that may meet their immediate needs, it does not meet the needs of users. The reason the Web is so successful is that services such as Yahoo, Lycos and Excite cut across the entire network of networks. The success of those portals is testimony to the existence of this need for global access points.

Waiting for a worldwide standard to emerge from several competing proprietary systems may be risky. Solutions that strive towards interoperability and a certain degree of harmony should be offered by key players who seek out, analyze, and, where appropriate, help to develop those solutions. We need a forum where those solutions can be freely discussed and considered. It would seem that the WIPO Advisory Committee on Management of Copyright and Related Rights in Global Information Networks could identify such a proper forum, and perhaps act as that forum itself. One immediate job for that forum would be to establish a directory of directories that covers the various categories of works, manifestations, and products. A second job would be to gather information on the work being done on identifiers, metadata, and related activities to have current information all in one place.

Digital Is Inevitable

The content is there. In almost all cases, it is in digital form or can be digitized. Networks with sufficient bandwidth are being built, and many business users and individual consumers are already connected. They are ready for the content. Many copyright industries and other rights holders are coming to the view that global networks represent good business opportunities and that digital, though it may be different, is nonetheless interesting commercially. In fact, it may be the only future growth area. [20][#N20]To put it simply, digital is inevitable. If most of the ingredients for successful electronic copyright commerce are assembled, the real question is: what do we do now?

Unless we want all content to be available free without restrictions — which would destroy copyright — rights-management solutions must be found. Viable electronic copyright-management systems must provide both blanket and transactional solutions. Interoperability is vital. The answers have been proposed; now we need to implement them.

[/j/jep/images/3336451.0004.303-00000001.gif]

Daniel J. Gervais is director of international relations and acting director of rightsholder relations at the Copyright Clearance Center, Inc., the world's largest reproduction-rights organization. Prior to joining CCC, he had been head of copyright projects at the World Intellectual Property Organization (WIPO), legal officer at the GATT (World Trade Organization), directorgeneral of the International Video Federation, and assistant secretary general of the Paris-based International Confederation of Societies of Authors and Composers (CISAC). He is a consultant to the Paris-based Organization for Economic Cooperation and Development (OECD). A Canadian lawyer and member of the Montreal Bar, Mr. Gervais practiced intellectual property law at a Montreal law firm for five years. He holds a Doctor of Laws degree magna cum laude from Nantes University (France), a Diploma in International Copyright Law magna cum laude from the Graduate Institute of Advanced International Studies (Geneva), a Masters degree (LL.M.) in computer law from the University of Montreal, as well as a Bachelor of Laws degree from the same university. Dr. Gervais speaks French, English, Spanish and German. He is the author of many articles and book chapters on intellectual-property law and rights management, published in six different languages.

Notes

1. The expression Electronic Copyright Management System (ECMS) is widely used. However, related rights may also be administered in this way. A better expression would be "Electronic Rights Management System", but we opted for the most common terminology. [#N1-ptr1]

2. For a discussion of the various collective management organization models, see Mihaly Ficsor. Collective Administration of Copyright and Neighboring Rights. WIPO, 1990. [#N2-ptr1]

3. While this type of application is newer, theatrical performance of theater plays has functioned under this model for a very long time, but does not enter the scope of this paper which focuses on diffusion techniques, i.e., on reception of material by users other than by direct personal access (presence at a live concert, etc.). [#N3-ptr1]

4. The adoption of international exhaustion (a territory of exhaustion might be the European Union) may impact on the application of this principle, but the principle remains nonetheless. [#N4-ptr1]

5. See for example the on-line contract for Mira uses, Article 3, http://www.mira.com/my/HelpCenter/Agreements.aspx?msid=2800 [http://www.mira.com/my/HelpCenter/Agreements.aspx?msid=2800]. [#N5-ptr1]

6. See Daniel Gervais, "The Law and Practice of Digital Encryption", op. cit. [#N6-ptr1]

7. Brian Green and Mark Bide. Unique Identifiers: a brief introduction. http://www.bic.org.uk/uniquid.html. [http://www.bic.org.uk/uniquid.html] IEEE has published a number of papers on ECMS. Among the most interesting ones, see A.K. Choudhury, N.F. Maxemchuk, S. Paul, and H.G. Schulzrinne. "Copyright Protection for Electronic Publishing Over Computer Networks". *IEEE Network,* Vol 9, Iss 3, p 12-20, May-June 1995. [#N7-ptr1]

8. ANSI/NISO standard Z39.56. [#N8-ptr1]

9. The Digital Object Identifier Initiative: Current Position and View Forward. Version 3, August 1998. http://www.doi.org/white-paper-3.pdf [http://www.doi.org/white-paper-3.pdf], at p. 10. [#N9-ptr1]

10. From a discussion on the DOI discussion group, http://www.doi.org/mailman/listinfo/discuss-doi [http://www.doi.org/mailman/listinfo/discuss-doi] [#N10-ptr1]

11. Document ISO/TC 46/SC 9/WG 1 N 64, dated September 24, 1998. [#N11-ptr1]

12. "3.1 audiovisual work: Work consisting of a series of related images, with or without accompanying sound, which is intended to be made visible and/or audible through the use of devices, regardless of the medium of initial or subsequent fixation." [#N12-ptr1]

13. Stuart Weibel, Erik Jul and Keith Shafer. PURLs: Persistent Uniform Resource Locators. 12/9/1998. Available at http://purl.oclc.org/oclc/purl/summary [http://purl.oclc.org/oclc/purl/summary] [#N13-ptr1]

14. Participants came from the following countries: Australia, Canada, Denmark, Finland, France, Germany, Japan, Norway, Sweden, Thailand, United Kingdom and the United States. [#N14-ptr1]

15. HTML or Hypertext Markup Language is a form of SGML, the Standard Generalized Markup Language, which is ISO standard 8879. SGML could be viewed as a system that allows one to define and use documents, in particular on the World Wide Web. XML is the Extensible Markup

Language. XML is a more "transaction-friendly" SGML-based system. It is thus probably better suited to enable e-commerce. —For a history, see http://www.xml.com/xml/ pub/w3j/s1.discussion.html [http://www.xml.com/] ⬆ [#N15-ptr1]

16. The National Library of the Netherlands is quite active in this field. It is also building a repertory of all online resources available in the Netherlands. See http://www.kb.nl/coop/donor/ [http://www.kb.nl/hrd/dd/dd_projecten/projecten_nedlib-en.html] ⬆ [#N16-ptr1]

17. For the purposes of this section, "privacy" relates to protection of consumer data, while "confidentiality" applies in a corporate environment. ⬆ [#N17-ptr1]

18. Anne W. Branscomb. Anonymity, Autonomy and Accountability: Challenges to the First Amendment in Cyberspaces. (1995) 104 Yale Law Journal, 1639. Professor Julie Cohen (University of Pittsburgh) has taken a similar view. ⬆ [#N18-ptr1]

19. Online full assignments (transfers) of rights may take place, but they would seem to be the exception. ⬆ [#N19-ptr1]

20. It must be said that some "newer" rights holders who started to play that role due to the possibilities of digital technology have (impatiently) held that view since the very beginning. ⬆ [#N20-ptr1]

Links from this article

Association of American Publishers — http://www.publishers.org [http://www.publishers.org]

The BIBLINK Project — http://hosted.ukoln.ac.uk/biblink/ [http://hosted.ukoln.ac.uk/biblink/]

ByLine: The Global Online Journalism Bank — [formerly http://www.universalbyline.com/scoop.html]

CNRI Handle System® — http://www.handle.net/index.html [http://www.handle.net/index.html]

Copyright Clearance Center — http://www.copyright.com/ [http://www.copyright.com/]

Copyright Licensing Agency - CLARCS — http://www.cla.co.uk/www/clarcs.html [http://www.cla.co.uk/support/clarcs/index.html]

Corporation for National Research Initiatives — http://www.cnri.reston.va.us /about_cnri.html [http://www.cnri.reston.va.us/about_cnri.html]

The Digital Object Identifier Initiative: Current Position and View Forward — http://www.doi.org/white-paper-3.pdf [http://www.doi.org/white-paper-3.pdf]

Discuss-DOI — http://www.doi.org/mailman/listinfo/discuss-doi [http://www.doi.org/mailman/listinfo/discuss-doi]

DONOR Homepage — [formerly http://www.kb.nl/coop/donor/]

Dublin Core Metadata Initiative — http://purl.oclc.org/dc/ [http://purl.oclc.org/dc/]

The IMPRIMATEUR Business Model V2.1 — [formerly http://www.imprimatur.alcs.co.uk/IMP_FTP/BMv2.pdf]

Inc. Republication Request Form — http://www.copyright.com/default.asp
[http://www.copyright.com/default.asp]

International DOI Foundation — http://www.doi.org/DOI-Found-Recruit.html
[http://www.doi.org/DOI-Found-Recruit.html]

International Federation of Reproduction Rights Organisations — http://www.ifrro.org/
[http://www.ifrro.org/]

Media Image Resource Alliance — http://www.mira.com/ [http://www.mira.com/]

Metadata Activity — http://www.w3.org/Metadata/Activity.html
[http://www.w3.org/Metadata/Activity.html]

MIRA Terms and Conditions — http://www.mira.com/Services/MoreTermsConditions.htm
[http://www.mira.com/Services/MoreTermsConditions.htm]

Stuart Weibel, Erik Jul and Keith Shafer. "PURLs: Persistent Uniform Resource Locators" —
http://purl.oclc.org/oclc/purl/summary [http://purl.oclc.org/oclc/purl/summary]

The Road to XML: Adapting SGML to the Web — [formerly
http://www.xml.com/xml/pub/w3j/s1.discussion.html

Stanford Digital Library Metadata Architecture —
http://citeseer.nj.nec.com/baldonado97stanford.html
[http://dbpubs.stanford.edu:8091/diglib/pub/delos.html]

Brian Green and Mark Bide. "Unique Identifiers: a brief introduction" —
http://www.bic.org.uk/uniquid.html [http://www.bic.org.uk/uniquid.html]

U.S. Fair Use Guidelines — http://lcweb.loc.gov/copyright/circs/circ21.pdf
[http://lcweb.loc.gov/copyright/circs/circ21.pdf]

U.S. MARC Database — http://lcweb.loc.gov/marc/ [http://lcweb.loc.gov/marc/]

Web Naming and Addressing Overview (URIs, URLs, . . .) — http://www.w3.org/Addressing/
[http://www.w3.org/Addressing/]